Recall:

Proposition: S is positive definite (resp. positive semi-definite) if and only if $x^TSx > 0 \ \forall x \neq 0$ (resp. $x^TSx \geq 0 \ \forall x$)

Proof:

($\Leftarrow$) Let $Sx = \lambda x$, where $x \neq 0$.

Then $0 < x^TSx = x^T(\lambda x) = \lambda x^Tx = \lambda \|x\|$ (resp. $0 \leq \lambda \|x\|$)

Since $\|x\| > 0$, we have $\lambda > 0$. (resp. $\lambda \geq 0$)

($\Rightarrow$) Let x be any nonzero vector

Since S is symmetric, by the Spectral theorem,

S has n orthonormal eigenvectors $v_1, \dots, v_n$

Hence $\{v_1, \dots, v_n\}$ is an orthonormal basis for $\mathbb{R}^n$

Let $\lambda_1, \dots, \lambda_n$ be the corresp. eigenvalues.

Since S is positive-definite, $\lambda_i > 0 \ \forall i$.

Now $x = c_1v_1 + \cdots + c_nv_n$ for some weights $c_1, \dots, c_n$ not all 0

$$\Rightarrow Sx = c_1Sv_1 + \cdots + c_nSv_n = c_1\lambda_1 v_1 + \cdots + c_n\lambda_n v_n$$

$$\Rightarrow x^TSx = (c_1v_1^T + \cdots + c_nv_n^T)(c_1\lambda_1 v_1 + \cdots + c_n\lambda_n v_n)$$

$$= c_1^2\lambda_1 + \cdots c_n^2\lambda_n > 0 \quad \text{since } \lambda_i > 0 \ \forall i.$$

(resp. ≥ 0 since $\lambda_i \geq 0 \ \forall i$) ∎

Singular Value Decomposition (SVD)

Let A be any $m \times n$ matrix.

A singular value decomposition of A is:

$$\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V^T}$$

where U and V are orthogonal matrices and Σ is a matrix whose only nonzero entries are on its "diagonal" and appear in decreasing order.

This decomposition has various applications including image processing/compression, best fit subspaces, and principle component analysis (towards Machine learning) Before we can discuss any of these, we need to understand SVD.

Ex: $\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2/\sqrt{6} & 0 & -1/\sqrt{3} \\ -1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$

Ex: $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \approx \begin{pmatrix} .763 & -.646 \\ .646 & .763 \end{pmatrix} \begin{pmatrix} 3.54 & 0 \\ 0 & 2.54 \end{pmatrix} \begin{pmatrix} -.646 & .763 \\ -.763 & .646 \end{pmatrix}$

Note: $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$ is not diagonalizable. So SVD $\neq$ Diagonalization in general. However, sometimes they do coincide (recitation)

We will see that: if $A = U\Sigma V^T$ has rank r, then

$$U = [u_1 \cdots u_m], \quad V = [v_1 \cdots v_n], \quad \Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{bmatrix}$$

$u_1 \cdots u_m$, $v_1 \cdots v_n$: singular vectors

$\sigma_1, \ldots, \sigma_r$: singular values

where:

1.) $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ and $\sigma_i = \sqrt{\lambda_i}$, where $\lambda_i \neq 0$ is an eigenvalue of A^TA (and AA^T)

2.) $\{v_1, \ldots, v_n\}$ are orthonormal eigenvectors of A^TA and:

- $\{v_1, \ldots, v_r\}$ be an orthonormal basis of $C(A^T)$
- $\{v_{r+1}, \ldots, v_n\}$ be an orthonormal basis of $N(A)$

3.) $\{u_1, \ldots, u_m\}$ are orthonormal eigenvectors of AA^T and

- $\{u_1, \ldots, u_r\}$ be an orthonormal basis of $C(A)$
- $\{u_{r+1}, \ldots, u_m\}$ be an orthonormal basis of $N(A^T)$

Ex:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

$r = 2$

Basis for $C(A)$

Basis for $N(A^T)$

2 singular values decreasing order

Basis for $C(A^T)$

($N(A) = \{0\}$ $\Rightarrow$ No basis).

How do we find an SVD?

Let A be $m \times n$ of rank r.

To find an SVD $A = U\Sigma V^T$ for A:

1. Orthogonally diagonalize $A^TA = QDQ^T$ with eigenvalues placed in D in decreasing order

2. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ be the eigenvalues of A^TA

   Set $\Sigma = \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \ddots & & O \\ & & \sqrt{\lambda_r} & \\ O & & & \end{bmatrix}$ and $V = Q$

3. Set $u_i = \frac{Av_i}{\sqrt{\lambda_i}}$ $\forall\ 1 \leq i \leq r$ where $v_i = i^{th}$ column of V.

   Let $\{u_{r+1}, \ldots, u_m\}$ be an orthonormal basis of $N(A^T)$

   Set $U = [u_1 \cdots u_m]$.

Ex: $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

1. $A^TA = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ has orthog. diag $\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}^{-1}$ (Check)

2. $\Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$

3. $u_1 = \frac{1}{\sqrt{3}} A \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} 2/\sqrt{2} \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$, $u_2 = \frac{1}{1} A \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

   Basis for $N(A^T)$ is $\left\{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\right\}$ (Check)

$$\Rightarrow u_3 = \begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \Rightarrow U = \begin{bmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$$